

0590
1/24

OIPE

#3

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:38

Input Set : N:\Crf3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

SEQUENCE LISTING

C--> 3 (1) GENERAL INFORMATION:

C--> 5 (i) APPLICANT: Stefan Somlo and Toshio Mochizuki

7 (ii) TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE
8 AND USES THEREOF

10 (iii) NUMBER OF SEQUENCES: 12

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN

14 (B) STREET: 90 PARK AVENUE

15 (C) CITY: NEW YORK

16 (D) STATE: NEW YORK

17 (E) COUNTRY: U.S.A.

18 (F) ZIP: 10016

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE
22 DISKETTE

23 (B) COMPUTER: IBM PC COMPATIBLE

24 (C) OPERATING SYSTEM: MS-DOS

25 (D) SOFTWARE: ASCII

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/753,008

C--> 29 (B) FILING DATE: 02-Jan-2001

35 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 09/385,752

33 (B) FILING DATE: 1999-08-30

36 (A) APPLICATION NUMBER: 08/651,999

37 (B) FILING DATE: MAY 23, 1996

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: ELIZABETH A. BOGOSIAN

41 (B) REGISTRATION NUMBER: 39,911

42 (C) REFERENCE/DOCKET NUMBER: 96700/395

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: (212) 697-5995

46 (B) TELEFAX: (212) 286-0854 or 286-0082

47 (C) TELEX: TWX 710-581-4766

50 (2) INFORMATION FOR SEQ ID NO: 1

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 866 amino acids

54 (B) TYPE: amino acid

55 (C) STRANDEDNESS: single

56 (D) TOPOLOGY: linear

W--> 58 (ii) MOLECULE TYPE:

59 (A) DESCRIPTION: peptide

ENTERED

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:38

Input Set : N:\Crf3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

```

61 (iii) HYPOTHETICAL: NO
63 (v) FRAGMENT TYPE: internal fragment
65 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1
67 Met Val Asn Ser Ser Arg Val Gln Pro Gln
68 1 5 10
70 Gln Pro Gly Asp Ala Lys Arg Pro Pro Ala
71 15 20
73 Pro Arg Ala Pro Asp Pro Gly Arg Leu Met
74 25 30
76 Ala Gly Cys Ala Ala Val Gly Ala Ser Leu
77 35 40
79 Ala Ala Pro Gly Gly Leu Cys Glu Gln Arg
80 45 50
82 Gly Leu Glu Ile Glu Met Gln Arg Ile Arg
83 55 60
85 Gln Ala Ala Ala Arg Asp Pro Pro Ala Gly
86 65 70
88 Ala Ala Ala Ser Pro Ser Pro Pro Leu Ser
89 75 80
91 Ser Cys Ser Arg Gln Ala Trp Ser Arg Asp
92 85 90
94 Asn Pro Gly Glu Glu Glu Ala Glu Glu Glu
95 95 100
97 Glu Glu Glu Val Glu Gly Glu Glu Gly Gly
98 105 110
100 Met Val Val Glu Met Asp Val Glu Trp Arg
101 115 120
104 Pro Gly Ser Arg Arg Ser Ala Ala Ser Ser
105 125 130
107 Ala Val Ser Ser Val Gly Ala Arg Ser Arg
108 135 140
110 Gly Leu Gly Gly Tyr His Gly Ala Gly His
111 145 150
113 Pro Ser Gly Arg Arg Arg Arg Arg Glu Asp
114 155 160
116 Gln Gly Pro Pro Cys Pro Ser Pro Val Gly
117 165 170
119 Gly Gly Asp Pro Leu His Arg His Leu Pro
120 175 180
122 Leu Glu Gly Gln Pro Pro Arg Val Ala Trp
123 185 190
125 Ala Glu Arg Leu Val Arg Gly Leu Arg Gly
126 195 200
128 Leu Trp Gly Thr Arg Leu Met Glu Glu Ser
129 205 210
131 Ser Thr Asn Arg Glu Lys Tyr Leu Lys Ser
132 215 220
134 Val Leu Arg Glu Leu Val Thr Tyr Leu Leu
135 225 230

```

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:38

Input Set : N:\Crf3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

137	Phe	Leu	Ile	Val	Leu	Cys	Ile	Leu	Thr	Tyr	
138					235					240	
140	Gly	Thr	Glu	Ala	Asp	Asn	Arg	Ser	Phe	Ile	
141					245					250	
143	Phe	Tyr	Glu	Asn	Leu	Leu	Leu	Gly	Val	Pro	
144					255					260	
146	Arg	Ile	Arg	Gln	Leu	Arg	Val	Arg	Asn	Gly	
147					265					270	
149	Ser	Cys	Ser	Ile	Pro	Gln	Asp	Leu	Arg	Asp	
150					275					280	
152	Glu	Ile	Lys	Glu	Cys	Tyr	Asp	Val	Tyr	Glu	
153					285					290	
155	Thr	Ala	Ala	Gln	Val	Ala	Ser	Leu	Lys	Lys	
156					295					300	
158	Asn	Val	Trp	Leu	Asp	Arg	Gly	Thr	Arg	Ala	
159					305					310	
161	Thr	Phe	Ile	Asp	Phe	Ser	Val	Tyr	Asn	Ala	
162					315					320	
164	Asn	Ile	Asn	Leu	Phe	Cys	Val	Val	Arg	Leu	
165					325					330	
167	Leu	Val	Glu	Phe	Pro	Ala	Thr	Gly	Gly	Val	
168					335					340	
170	Ile	Pro	Ser	Trp	Gln	Phe	Gln	Pro	Leu	Lys	
171					345					350	
173	Leu	Ile	Arg	Tyr	Val	Thr	Thr	Phe	Asp	Phe	
174					355					360	
176	Phe	Leu	Ala	Ala	Cys	Glu	Ile	Ile	Phe	Cys	
177					365					370	
179	Phe	Phe	Ile	Phe	Tyr	Tyr	Val	Val	Glu	Glu	
180					375					380	
W--> 182	Ile	Leu	Glu	Xaa	Ile	Arg	Ile	His	Lys	Leu	
183					385					390	
W--> 185	His	Tyr	Phe	Arg	Xaa	Ser	Phe	Trp	Asn	Cys	
186					395					400	
188	Leu	Asp	Val	Val	Ile	Val	Val	Leu	Ser	Val	
189					405					410	
191	Val	Ala	Ile	Gly	Ile	Asn	Ile	Tyr	Arg	Thr	
192					415					420	
W--> 194	Ser	Asn	Val	Glu	Val	Xaa	Leu	Leu	Gln	Phe	
195					425					430	
W--> 197	Leu	Xaa	Glu	Asp	Gln	Asn	Thr	Phe	Pro	Asn	
198					435					440	
200	Phe	Glu	His	Leu	Ala	Tyr	Trp	Gln	Ile	Gln	
201					445					450	
203	Phe	Asn	Asn	Ile	Ala	Ala	Val	Thr	Val	Phe	
204					455					460	
206	Phe	Val	Trp	Ile	Lys	Leu	Phe	Lys	Phe	Ile	
207					465					470	
209	Asn	Phe	Asn	Arg	Thr	Met	Ser	Gln	Leu	Ser	

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:38

Input Set : N:\Crf3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

210		475	480
212	Thr Thr Met Ser Arg Cys Ala Lys Asp Leu		
213		485	490
215	Phe Gly Phe Ala Ile Met Phe Phe Ile Ile		
216		495	500
218	Phe Leu Ala Tyr Ala Gln Leu Ala Tyr Leu		
219		505	510
221	Val Phe Gly Thr Gln Val Asp Asp Phe Ser		
222		515	520
224	Thr Phe Gln Glu Cys Ile Phe Thr Gln Phe		
225		525	530
227	Arg Ile Ile Leu Gly Asp Ile Asn Phe Ala		
228		535	540
W--> 230	Glu Ile Glu Glu Ala Asn Xaa Arg Val Leu		
231		545	550
233	Gly Pro Ile Tyr Phe Thr Thr Phe Val Phe		
234		555	560
236	Phe Met Phe Phe Ile Leu Leu Asn Met Phe		
237		565	570
239	Leu Ala Ile Ile Asn Asp Thr Tyr Ser Glu		
240		575	580
W--> 242	Val Lys Ser Asp Leu Xaa Xaa Xaa Ala Gln		
243		585	590
245	Gln Lys Ala Glu Met Glu Leu Ser Asp Leu		
246		595	600
248	Ile Arg Lys Gly Tyr His Lys Ala Leu Val		
249		605	610
251	Lys Leu Lys Leu Lys Lys Asn Thr Val Asp		
252		615	620
254	Asp Ile Ser Glu Ser Leu Arg Gln Gly Gly		
255		625	630
257	Gly Lys Leu Asn Phe Asp Glu Leu Arg Gln		
258		635	640
260	Asp Leu Lys Gly Lys Gly His Thr Asp Ala		
261		645	650
263	Glu Ile Glu Ala Ile Phe Thr Lys Tyr Asp		
264		655	660
266	Gln Asp Gly Asp Gln Glu Leu Thr Glu His		
267		665	670
269	Glu His Gln Gln Met Arg Asp Asp Leu Glu		
270		675	680
272	Lys Glu Arg Glu Asp Leu Asp Leu Asp His		
273		685	690
275	Ser Ser Leu Pro Arg Pro Met Ser Ser Arg		
276		695	700
278	Ser Phe Pro Arg Ser Leu Asp Asp Ser Glu		
279		705	710
281	Glu Asp Asp Asp Glu Asp Ser Gly His Ser		
282		715	720

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:38

Input Set : N:\Crif3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

```

284      Ser Arg Arg Arg Gly Ser Ile Ser Ser Gly
285                      725                      730
287      Val Ser Tyr Glu Glu Phe Gln Val Leu Val
288                      735                      740
290      Arg Arg Val Asp Arg Met Glu His Ser Ile
291                      745                      750
293      Gly Ser Ile Val Ser Lys Ile Asp Ala Val
294                      755                      760
296      Ile Val Lys Leu Glu Ile Met Glu Arg Ala
297                      765                      770
299      Lys Leu Lys Arg Arg Glu Val Leu Gly Arg
300                      775                      780
302      Leu Leu Asp Gly Val Ala Glu Asp Glu Arg
303                      785                      790
305      Leu Gly Arg Asp Ser Glu Ile His Arg Glu
306                      795                      800
308      Gln Met Glu Arg Leu Val Arg Glu Glu Leu
309                      805                      810
311      Glu Arg Trp Glu Ser Asp Asp Ala Ala Ser
312                      815                      820
314      Gln Ile Ser His Gly Leu Gly Thr Pro Val
315                      825                      830
317      Gly Leu Asn Gly Gln Pro Arg Pro Arg Ser
318                      835                      840
W--> 320      Ser Arg Pro Ser Ser Ser Gln Ser Xaa Thr
321                      845                      850
323      Glu Gly Met Glu Gly Ala Gly Gly Asn Gly
324                      855                      860
326      Ser Ser Asn Val His Val
327                      865
C--> 330 (2) INFORMATION FOR SEQ ID NO: 2
332      (i) SEQUENCE CHARACTERISTICS:
333          (A) LENGTH: 523 amino acids
334          (B) TYPE: amino acid
335          (C) STRANDEDNESS: single
336          (D) TOPOLOGY: linear
W--> 338      (ii) MOLECULE TYPE:
339          (A) DESCRIPTION: peptide
341      (iii) HYPOTHETICAL: NO
343      (v) FRAGMENT TYPE: internal fragment
345      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2
348          Phe Leu Ala Lys Glu Glu Ala Arg Lys Val
349              1              5              10
351          Lys Arg Leu His Gly Met Leu Arg Ser Leu
352              15              20
354          Leu Val Tyr Met Leu Phe Leu Leu Val Thr
355              25              30
357          Leu Leu Ala Ser Tyr Gly Asp Ala Ser Cys
358              35              40

```

VERIFICATION SUMMARY

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:39

Input Set : N:\Crif3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

L:3 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:5 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]
L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:58 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:182 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:194 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:197 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:230 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:330 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:338 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=2
L:360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:363 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:375 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:378 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:381 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:397 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:448 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:478 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:508 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:516 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:544 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:647 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:655 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4
L:671 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:683 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:689 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:713 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:716 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:728 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:731 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:734 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:737 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:740 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:743 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:746 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:776 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:784 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:801 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]

VERIFICATION SUMMARY

DATE: 02/11/2002

PATENT APPLICATION: US/09/753,008

TIME: 09:36:39

Input Set : N:\Crf3\RULE60\09753008.txt

Output Set: N:\CRF3\02112002\I753008.raw

L:809 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:921 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:929 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7
L:1121 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1129 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8
L:1141 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1149 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:1161 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1169 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:1182 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1190 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:1201 M:220 C: Keyword misspelled or invalid format, [(2) INFORMATION FOR SEQ ID NO:]
L:1209 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12